

## Geography

### 1.0 Geography Division Overview

Although the Geography Division is part of the Decennial program area, we discuss it separately in our Plan as a special interest item. This section describes two major components: the Geographic Support System (GSS), our IT support for most of the U.S. Census Bureau's censuses and surveys; and the GSS Infrastructure, which comprises our Division's infrastructure.

The Geographic Support System is an integrated and automated computer-based system used to establish and maintain accurate geographic boundaries within the Topologically Integrated Geographic Encoding and Referencing (TIGER ) System and accurate addresses in the Master Address File.

The GSS Infrastructure includes IT resources that are not program specific but are part of the support structure for the day-to-day activities of our employees. This includes our office automation and Local Area Network support.

The goal of the Geography Division is to provide the basic maps, reference files, and associated processing systems needed to meet the geographic requirements of all U.S. Census Bureau programs. It has five basic components:

**Building and maintaining the TIGER database and the Master Address File of housing units and economic establishments:** the Master

Address File is crucial to conducting many U.S. Census Bureau programs, including the 2000 Decennial Census and the American Community Survey. Continuously updating the Master Address File is the most cost-effective and quality-assured method for providing a housing unit address list.

**Continuing to link the Master Address File with the TIGER database:** this crucial on-going effort involves linking the Master Address File to the TIGER database, which produces maps and the names of all governmental units for data tabulations. TIGER also allows us to assign every housing unit and business establishment address to the correct census block for processing and tabulation. Linking the Master Address File to the TIGER database is the most efficient and cost effective way of adding new streets and housing units and their locations and ZIP codes to U.S. Census Bureau databases.

**Continuing our partnerships with state, local, and tribal agencies:** we are continuing to support partnerships with state, local, and tribal agencies. To address past concerns expressed by officials at all levels of government and to increase the confidence of our customers with census results, we provide opportunities for local officials to review existing information about streets, boundaries, and addresses. We listen to suggestions from those officials

## 1.1 Geography Division Products, Services, and Customers

The Geography Division provides the following major products and services to these customers:

Products/Services	Customers
paper maps* digital maps* TIGER/Line TIGER/Census Tract Street Index files TIGER/Census Tract Comparability files	all data users and U.S. Census Bureau program areas
TIGER database Master Address File	U.S. Census Bureau program areas only

\*These products are also provided to the U.S. Census Bureau's private sector partners who help keep the spatial database up-to-date. Samples of these products are available on the Internet at [www.census.gov/geo/www/index.html](http://www.census.gov/geo/www/index.html).

Our customers often repackage our data and offer them on CD-ROM, the Internet, or other media. These customers depend on us to continuously maintain our Geographic Support System. A large number of firms (you can find a list of these vendors at [www.census.gov/geo/www/tiger/vendors/html](http://www.census.gov/geo/www/tiger/vendors/html)) use various U.S. Census Bureau extract data/programs for the following purposes:

- address matching;
- geocoding;
- mapping;
- redistricting services;
- vehicle routing and dispatching;
- desktop mapping;
- land survey;
- site location;
- marketing analysis and studies; and
- transportation management.

Examples of public and private sector applications that use TIGER extract products include:

- the New York City Metropolitan Transportation Authority, for devising more efficient routes for carpooling, snow-clearing, meter-reading, and garbage-collecting in their 32 counties;
- the Newport News, Virginia, waterworks, to extract products and its own water line database to create a geographic information system that includes a 250 square mile territory;
- the Baltimore County, Maryland, Police Department, to geocode spousal abuse cases and link them with variables such as income and employment;
- the Maryland Department of State Planning, to produce customized files for each of Maryland's 23 counties and the City of Baltimore;
- local planners, to determine where services are needed for the elderly, child care facilities, recreation sites, schools, and libraries;
- ALK Associates of Princeton, New Jersey, to produce software used for routing;
- Caliper Corporation of Newton, Massachusetts, to produce several types of Geographic Information System software; and
- Chicago Itap Corporation of Lemont, Illinois, to produce a desktop street-mapping package.

In addition to the programs listed, the Geography Division is proposing a 21<sup>st</sup> Century Master Address File/TIGER new initiative. This initiative will allow the U.S. Census Bureau to further develop the capabilities of the Master Address File/TIGER system. This initiative also allows the sponsoring and

participating organizations to begin activities that will further develop the Master Address File/TIGER system by using new technology, commercial off-the-shelf software, and new address and geographic data sources. The IT-related costs for this initiative are reflected in the Census Modernization initiative.

### Supporting a “Digital” Department of Commerce

The following are just a few examples of how the Geography Division uses “digital” technology to support the Decennial Census and other U.S. Census Bureau activities:

**Intranet Geography Web System:** this is the primary means of providing information to Regional Offices and other clients within the U.S. Census Bureau.

**Map Production System:** map production is done at the Regional Census Centers, the National Processing Center, and Headquarters. The maps (in digital form as a Map Image Metafile) are then transmitted to the actual map plotting sites. Software then converts the Map Image Metafiles to output device-specific format (Hewlett-Packard Graphical Language or PostScript) and sends the resulting print file to the output device. Thus, map production is a two-step process. First, on a computer system where a Master Address File and TIGER database is located, a Map Image Metafile is created (the Map Image Metafile may contain more than one map sheet). The Map Image Metafile may then be processed for printing immediately, copied by network to another computer system for printing later, or copied to CD-Recordable for

shipment to a remote printing site (for example, Local Census Offices).

The Master Address File is a database of all addresses in the United States; the residential addresses in this database will be validated prior to Census 2000 mailing and enumeration activities. It is an essential component of every census operation in which housing unit addresses are used. For example, addresses in the Master Address File are critical to the production of maps. Thus, the Master Address File can be thought of an address database with geographic components that complements the TIGER database, a geographic database with address components. Programs and projects that rely on detailed address information use the Master Address File.

**Topologically Integrated Geographic Encoding and Referencing (TIGER®) System:** the TIGER® database provides geographic products such as maps and the spatial characteristics of all governmental units for data tabulations. It also allows us to automatically and electronically assign every residential and business establishment address to the correct spatial location for processing and tabulation.

- maps are produced at Regional Census Centers/the National Processing Center and are distributed locally;
- maps are produced at Regional Census Centers and sent to Local Census Offices for distribution;
- Map Image Metafiles are transferred onto CD-ROMs and forwarded to Local Census Offices for production and distribution.

For more specialized maps required for reports, and after “one-time” products, the U.S. Census Bureau does basic interactive map editing and map image manipulation using its graphic workstations and file servers as the host for the map files. This system uses Map Image Metafiles or the TIGER/Line files. During this process, we use both commercial and customized software for the graphic workstations and file servers.

### **Intranet Geography Web System**

The Intranet Geography Web System is the primary means of providing information to Regional Offices and other clients

within the U.S. Census Bureau. Available information includes:

- GSS Help System (program and clerical reference manuals);
- problem referrals;
- Geographic Update System documentation;
- TIGER documentation;
- production status reports; and
- Geographic IT configuration, support, and documentation.

### **Enterprise Support**

The Geographic Update System, the Map Production System, and the Geography Intranet Web System are managed and supported within the Geography Division. The sole exception is the OpenVMS Alpha cluster of servers located in the Bowie Computer Center. All of the large-scale UNIX servers and UNIX workstations in the Regional Census Centers, the National Processing Center, and Headquarters are part of the enterprise support performed under the Geographic Support System.

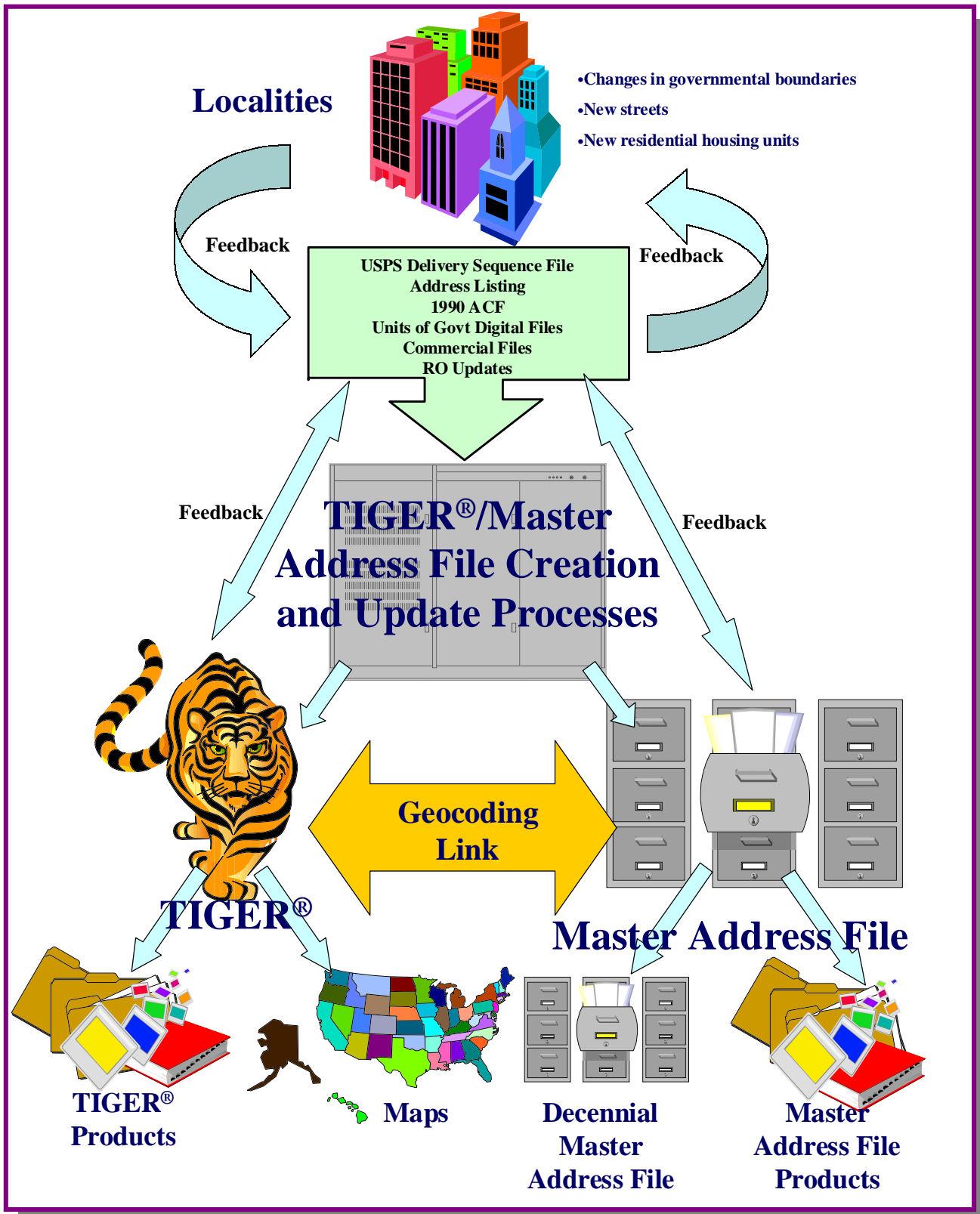


Figure 1: Basic Interactive Data Flow Associated with the Geographic Support System and Updating TIGER and the Master Address File

the Boundary and Annexation Survey incorporated places with a population of 2500+. In years ending in 8,9, and 0, the Decennial Census helps us expand the Boundary and Annexation Survey to include all legal entities, regardless of size.

The table opposite summarizes the different entities included in the Boundary and Annexation Survey:

Years	All Sizes	Entities 2500+	Entities 5000+
ending in 1,3,4,5,6			✓
ending in 2, 7		✓	✓
ending in 8,9,0	✓	✓	✓

## Decennial Support Programs

### Census 2000 Redistricting Data Program

This program encompasses the various activities that create the geographic context for tabulating Census 2000 data used by the states in their redistricting activities under Public Law 94-171. During the Block Boundary Suggestion Project, the U.S. Census Bureau offers state redistricting officials the opportunity to incorporate features they request be held as Census 2000 block boundaries.

During the Voting District Project, the state redistricting officials will submit the boundaries of voting districts (precincts) and state legislative districts that follow many of the features they provided during the Block Boundary Suggestion Project.

### Participant Statistical Areas Programs

These programs enable local participants to delineate, following U.S. Census Bureau guidelines, statistical areas such as census tracts, block groups, census

designated places, and census county divisions.

These statistical areas are the basis for the tabulation of Census 2000 data and are used extensively by both the public and private sectors.

### Urbanized Area Delineation

The U.S. Census Bureau delineates urbanized areas to provide a better separation of urban and rural territory, population, and housing in the vicinity of larger places.

The 1990 definition of an urbanized area is an area consisting of a central place(s) and adjacent urban fringe that together have a minimum residential population of at least 50,000 people and generally an overall population density of at least 1,000 people per square mile of land area. The U.S. Census Bureau uses published criteria to determine the qualifications and boundaries of urbanized areas.

The 2000 urbanized area definition (criteria) is in progress and will be completed in fall of 1999.



The computer clusters that the Geographic Support System uses are given in the following table:

Description	Location
Geo Bowie Computer Center Production Cluster	Bowie Computer Center
Geo HQ Development Cluster	HQ; Geo Division
Geo Distributive Production Cluster	HQ; National Processing Center; Regional Census Centers
Research Server	HQ; Geo Division

Each cluster has the following configuration:

#### Geography Bowie Computer Center Production Cluster

- multiple Compaq Computer (formerly Digital Equipment) Corporation computers grouped into an Open VMS Alpha cluster multi-processor environment with at least 2,000 GB of disk storage and 4 GB of memory; and
- Silicon Graphics UNIX file server with 16B of memory and 140 GB of disk storage for remote backup of geographic and decennial support data.

#### Geography Headquarters Development Cluster

- multiple UNIX multi-processor file servers with at least 200 GB disk storage for general software development. Associated with these servers are PCs and Silicon Graphics (SGI) graphic workstations;
- Compaq Computer (formerly Digital Equipment Corporation) Compaq Tru64 UNIX machine for software development;
- multiple Silicon Graphics UNIX multi-processor file servers with at least 200 GB disk storage (identical to the Regional Census Center configuration) for testing Regional Census Center applications;
- multiple Silicon Graphics UNIX multi-processor file servers, for map software development;
- Silicon Graphics UNIX multi-processor file server for production;
- Silicon Graphics UNIX multi-processor file server to research and develop new technologies (such as commercial off-the-shelf software to replace existing internally developed software).

- **E-Size color printers:** needed to maintain the TIGER /Master Address File databases through the expanded Boundary and Annexation Survey in 1999, Census 2000, and other geographic programs.
  - **Graphic workstations:** required to develop application software and to update and maintain the TIGER /Master Address Files databases.
  - **PCs:** to replace outdated and incompatible office automation equipment in Headquarters and to meet U.S. Census Bureau standards for use as an interface to TIGER /Master Address File databases by subject matter staff. Additional PCs are required to access the TIGER /Master Address File database in the Regional Census Centers and the National Processing Center for updating and maintaining addresses. We are using a three-year lifecycle for replacing the PCs. By July 1 of each year, we will determine which PCs will have priority for replacement in the following fiscal year.
- **UNIX multi-processor servers:** for use in the Regional Census Centers, National Processing Center, and Headquarters, for processing maps and other TIGER /Master Address File batch operations. Each server will have a large disk farm (minimum 100 GB) and upward scalability. Each will have a minimum of 1 GB of memory.
  - **UNIX servers:** we require dedicated servers to process local spatial data for modeling to the TIGER database and for supporting Oracle Relational Database Management System applications, Data Access and Dissemination System development, backup services and Intranet applications.
  - **Research server:** required to support Headquarters development activities and test new technologies.



<b>Geographic Support System Milestones, FY 00</b>					
<b>Description</b>	<b>Estimated</b>		<b>Actual</b>		<b>Progress to Date</b>
	<b>Start Date</b>	<b>Finish Date</b>	<b>Start Date</b>	<b>Finish Date</b>	
<b>Local Update of Census Address: Implement LUCA</b>	<b>07/97</b>	<b>10/99</b>			<b>In progress.</b>
<b>Geographic Services – Participant Statistical Entities: 1999 Boundary and Annexation Survey</b>	<b>06/98</b>	<b>11/99</b>			<b>In progress.</b>
<b>Build MAF</b>	<b>01/96</b>	<b>12/99</b>			<b>In progress.</b>
<b>Link City – Style MAF addresses to TIGER</b>	<b>01/95</b>	<b>01/00</b>			<b>In progress.</b>
<b>MAF Improvement of City-Style Address Areas: Targeted Map Updates</b>	<b>03/97</b>	<b>01/00</b>			<b>In progress.</b>
<b>Determine 2001 upgrade and replacement requirements for database and application servers</b>	<b>10/99</b>	<b>01/00</b>			
<b>Demographic Census Entities: Public Use Microdata Samples</b>	<b>06/97</b>	<b>02/00</b>			<b>In progress.</b>
<b>Traffic Analysis Zones</b>	<b>11/97</b>	<b>03/00</b>			<b>In progress.</b>
<b>Reduce number of network protocols to one: TCP/IP</b>	<b>04/99</b>	<b>03/00</b>			<b>In progress.</b>
<b>Reduce number of network operating systems to two: NFS, Common Internet Filing System</b>	<b>04/99</b>	<b>03/00</b>			<b>In progress.</b>
<b>Special Place/Group Quarters Inventory</b>	<b>04/96</b>	<b>04/00</b>			<b>In progress.</b>
<b>Geographic Services – Participant statistical Entities</b>	<b>06/96</b>	<b>04/00</b>			<b>In progress.</b>
<b>Local Update of Census Address: Implement Decennial Appeals Process</b>	<b>06/99</b>	<b>04/00</b>			<b>In progress.</b>
<b>Zip Code Tabulation Areas</b>	<b>04/97</b>	<b>08/00</b>			<b>In progress.</b>
<b>Final Boundary Validation Process</b>	<b>12/99</b>	<b>08/00</b>			
<b>Geographic Services – Participant Statistical Entities: 2000 Boundary and annexation survey</b>	<b>05/99</b>	<b>09/00</b>			<b>In progress.</b>

<b>Geographic Support System Milestones, FY 03</b>					
<b>Description</b>	<b>Estimated</b>		<b>Actual</b>		<b>Progress to Date</b>
	<b>Start Date</b>	<b>Finish Date</b>	<b>Start Date</b>	<b>Finish Date</b>	
<b>Congressional District Atlas</b>	<b>02/02</b>	<b>01/03</b>			
<b>Boundary Files – 108<sup>th</sup> Congressional Districts</b>	<b>09/02</b>	<b>01/03</b>			
<b>Congressional Districts (106<sup>th</sup>) and (108<sup>th</sup>)</b>	<b>02/98</b>	<b>04/03</b>			<b>In progress.</b>
<b>Demographic Census Entities: Metropolitan Areas</b>	<b>06/97</b>	<b>04/03</b>			<b>In progress.</b>
<b>Population and Housing Atlas</b>	<b>03/02</b>	<b>09/03</b>			

### 2.1.3 Geographic Support System Performance Measures

In last year's Operational IT Plan the Geography Division identified 11 performance goals for the Geographic Support System. However, we found these goals had two significant flaws: first, they were project goals rather than IT performance goals and, second, the statistics generated for these projects did not produce useful performance measures. The key measure was timely delivery of various geographic products, but

recording whether a deadline is met is not, in and of itself, a very useful performance metric. It does little to explain why a deadline was met or missed since there may be many factors involved. To be useful, performance metrics need to be more specific and focused. For these reasons, we replaced the original performance goals with the four new goals noted below:

<b>GSS Performance Measures</b>			
<b>Performance Goals</b>	<b>Performance Measures</b>	<b>Target Performance</b>	<b>Current Performance</b>
<b>Purchase hardware and software to support GSS programs</b>	<b>Average time to enter purchase request into system</b>	<b>Within 5 days</b>	<b>*</b>
<b>Distribute hardware and software</b>	<b>Time to deliver to end user</b>	<b>Within 5 days after receipt</b>	<b>*</b>
<b>Respond to Help Desk requests**</b>	<b>Average response time to resolve</b>	<b>Within 4 hours</b>	<b>*</b>
<b>Quality assurance of Geography products</b>	<b>Percent of products returned for rework</b>	<b>Less than 1%</b>	<b>*</b>

\* An automated system to establish baselines and measure performance is in progress.

\*\* Since our help desk supports both production and office automation IT resources, this performance goal is also included in section 3.3 Geographic Support System Infrastructure Performance Measures.

We are working with the Field and Decennial Directorates to create separate plans for each Regional Census Center; we are also working with Field to create a plan for the National Processing Center. Finally, we are creating a plan for Headquarters. Each of these plans is in progress and we expect to complete them by the end of 1999. All of these continuity plans are based on the distributed character of TIGER /Master Address File processing and on having duplicate hardware and software configurations at the Regional Census Centers, the National Processing Center, and Headquarters. Thus a processing disruption at any site will cause that

processing to be moved to another site until the disrupted site can be repaired.

We are also working on security plans for the Geographic Support System. Security plans, like the continuity plans, are being developed for each site where a portion of this system resides. We are working with the Security Office to determine the final aspects of the plans. Problems that still must be resolved include defining a computer system for purposes of the security plans, identifying the security standards that the U.S. Census Bureau uses, and deciding how these standards should be implemented in the Geographic Support System.

### 2.1.5 Geographic Support System References

The Geographic Support System is supported by the following planning documents:

- **Budget Submission for FY 2001, dated June 1999;**
- **1999 Strategic IT Plan, dated December 18, 1998, pages 77-81, 98;**
- ***Early Decennial Census Operations*, Requirement Initiative DC02-9702, approved October 3, 1997;**
- ***Large Format Printers*, Requirement Initiative DC01-9703, approved August 15, 1997;**
- ***Computer Hardware and Software*, Requirement Initiative DC01-9701, approved August 5, 1997;**
- ***Geography Workstations*, Requirement Initiative DC01-9505, approved August 11, 1995; and**
- ***A Geographic Support System for the U.S. Bureau of the Census: Functional Requirements for the System*, SPAD System, Ltd., Reston, VA, February 1983.**

from the WP1 building to the U.S. Census Bureau's Wide Area Network from 10 Mbps fiber to 155 Mbps Asynchronous Transfer Mode (ATM). The ATM link will include hardware encryption. At the same time, additional 100 Mbps switches will be installed to increase Local Area Network capacity. The Geography Division is funding these upgrades.

We currently run four network operating systems; these are:

- Network Filing System, usually to connect UNIX computers;
- Server Message Block, usually to connect Windows NT computers;
- Appleshare, usually to connect Macintosh computers; and
- Netware, usually to connect Windows computers.

Each of these network operating systems uses a different protocol on the network, namely TCP, NetBios, AppleTalk, and IPX. We are working with the IT Directorate to reduce the number of protocols to one, TCP/IP. We expect, within the next year, to also reduce our network operating systems to these:

- Network Filing System, usually to connect UNIX computers;
- Common Internet Filing System, usually to connect all other computers; and
- NetWare to connect legacy systems in the rest of the U.S. Census Bureau.

### 3.2 Geographic Support System Infrastructure Progress Against Planned Milestones

Geographic Support System Infrastructure Milestones, FY 98					
Description	Estimated		Actual		Progress to Date
	Start Date	Finish Date	Start Date	Finish Date	
No milestones to report					

Geographic Support System Infrastructure Milestones, FY 99					
Description	Estimated		Actual		Progress to Date
	Start Date	Finish Date	Start Date	Finish Date	
Obtain PC replacements for 2000	10/98	09/99	11/98		Purchased 50 PCs

Geographic Support System Infrastructure Milestones, FY 00					
Description	Estimated		Actual		Progress to Date
	Start Date	Finish Date	Start Date	Finish Date	
Determine PC replacement requirements for 2001	10/99	01/00			

GSS Performance Measures			
Performance Goals	Performance Measures	Target Performance	Current Performance
Respond to Help Desk requests	Average response time to resolve	Within 4 hours	Not yet implemented

### 3.4 Geographic Support System Infrastructure Risks

Hiring and retaining competent staff continues to be a central risk for the Geography Division infrastructure. A U.S. Census Bureau-sponsored job fair was held at Headquarters in late January 1999 and a large number of computer specialists (Series 0334, grades 5-12) were interviewed for many positions throughout the Geography Division. Since the job fair, we have not been able to hire any computer specialists.

A cursory review of the difficulty in hiring computer specialists in the Geography Division gives these observations:

- the government-wide general pay schedule is not competitive for computer specialists;
- for computer specialists, the U.S. Census Bureau participates in an Office of Personnel Management-sponsored merit selection process that is slow and error-prone;
- computer specialist candidates lacking up-to-date skills are not as valuable because of the rapid technology changes in the computer field.

### 3.5 Geographic Support System Infrastructure References

The Geographic Support System Infrastructure is supported by the following planning documents:

- *Enterprise Printer Resources*, Requirement Initiative PRMAP9701, approved May 13, 1998;
- *Macintosh Computers*, Requirement Initiative GE06-9901, approved August 5, 1997; and
- *Enterprise Personal Computer Management and Acquisition Plan*, Requirement Initiative PCMAP9601, approved December 5, 1995.